Hello everyone, my name is caoxi zhang. Today I’m going to present my final year project plan, traffic route planning for accident prediction.

This is the outline of today’s presentation. I will go through introduction, issues and solution, analysis methods of the results and the future work.

Now I’d like to introduce the main topic, machine learning in internet of things. In the modern society, there are lots of intelligent gadgets playing an important role in our life. Like google map gets the live traffic flow data from the sensors and illustrates them by color on your phone. This is the internet of things. And machine learning can be applied to it to make it even smarter. One of the applications I’m going to focus on is the smart transportation.

Before I start my project, I did a literature survey on serval mainstream smart transportation research. A number of recent studies focus on predicting short-term or long-term to make a more time efficient route optimization method. This is a table presenting the behavior of the algorithms in those studies. Then, I developed my idea from their research.

Let’s first look at some statics. Every 16 minutes, there is someone killed or injured on the road and in London, there were over 23 thousand accidents in 2021. The current issues are that most navigation apps on the market only focus on time efficiency rather than road safety.

So, my solution is traffic route planning for accident prediction. It used machine learning to make a safer and time efficient route suggestion.

This is my project route map. Next, I’ll present the data analysis of my dataset from the British government.

There are three major features in the dataset. The first one is the area. The left side is a heatmap showing the occurrence of accidents. It divides London city center into 100x100 blocks. The red means there are more accidents happened. The right-side graph shows the top five boroughs with the most accidents. And the number of accidents in different severities. 3 means slight. It is shown that certain areas having higher accidents rates.

Next two graphs show the relationships between the date and the time of accidents happened. It is clear to see that most of the accidents occurred on Friday and during the evening peak time around 6pm.

The last feature is the weather. Though most of the accidents happened on a sunny day, the special weather conditions can also have some affections, especially a rainy day.

Now let’s moving on to the machine learning parts. These are algorithms that can be used in this project. SVM, decision trees and BART. The main part of the project will be using classification. First training the machine learning models with the dataset. Then given a set of routes from start to end, crossing area a, b and c, the algorithm would classify each segment of the route into high, medium or low accident level. If one route crosses too many red areas, the machine would suggest the user other safer ways. It also considers the time and weather conditions, if it’s 5pm on Friday, there may be higher chances of collision in certain crowed areas. And if its rainy, there may be more accidents happened in rural areas.

After having the results from the models, accuracy function and confusion matrix would be used to evaluate them. Accuracy is the number of correct predictions divided by the total number of predictions. The confusion matrix visualizes the ground truth versus the predictions. TP and TN are the correct positive and negative results. FP and FN are the false positive and negative results. By combining the precision and the recall values using this table to get the f1 score, it will measure the behavior of the classification algorithms.

The last part is the future work. The next step is training the machine learning models and evaluating the results. Moreover, I’d like to include more features such as drivers and vehicle’s information to make the prediction more details. Also considering combining the regression algorithms predicting traffic flow including the speed limits with classification algorithms.

Thank you for listening, please feel free to ask me any question.